Algerian Journal of Environmental Science and Technology
March edition. Vol.9. N°1. (2023)
ISSN : 2437-1114
www.aljest.org

Critical Temperatures Prediction of Organic Compounds (Aliphatic Alkanes) Using QSPR Approach

B. Souyei[1*], N. Kertiou[2,3], A. Hadj Seyd[4], Y. Labbi[5], A. Khechekhouche[6]

[1]Chemistry Department, Echahid Hamma Lakhdar University of El Oued Algeria
[2] Faculty of Sciences and Technology - Khenchela, Abbes Laghrour University, Algeria
[3]Environmental and Food Safety Laboratory (08 EFSL), Badji Mokhtar -Annaba University, Algeria
[4]Faculty of Renewable Energy, Hydrocarbons, Earth and Universe Sciences, University of Kasdi Merbah, Algeria
[5]Electrical Engineering Department, University of El Oued Algeria
[6]Faculty of Technology, University of El Oued El-Oued, Algeria

*Corresponding author: belgacem3963_alg@yahoo.fr; Tel.: +213 0600000

ARTICLE INFO

Article History:

Received ***: 10/12/2020***
Accepted ***: 01/05/2021***

Key Words:
Aliphatic Alkanes;
Statistical Model;
hydrocarbon mixture;
MLR; PLS.

ABSTRACT/RESUME

Abstract: *The quantitative structure-property relationship QSPR method using Multiple Linear Regression MLR and Partial Least Squares PLS methodologies was performed for 160 organic compounds (hydrocarbons, branched alkanes, branched and unbranched alkenes, and alkynes). The MLR and PLS methods were employed to explore the correlation between the molecular descriptors which are the structural representation while the critical temperature T_c is the property representation. Using Dragon descriptors, this study was aimed at developing a predictive and robust QSPR models for predicting T_c. According to the squared correlation coefficients (R^2 =0.942 and 0.941), standard error (s =0.88 and 0.797) and the leave-one-out cross-validation correlation coefficients (Q^2_{Loo} = 0.834 and 0.932), for the MLR and PLS methods respectively, the results demonstrated almost identical qualities and good predictive ability for both the MLR and PLS models.*

I. Introduction

Estimating and determining the critical properties of pure compounds is of critical importance for finding out thermodynamic and volumetric properties when using corresponding state correlations. These parameters have a direct application in selecting the operating conditions of substances in gas or liquid phase, or in other aspects during production [1]. With regard to the different calculations of critical point, we find that, the calculations of critical temperatures and pressures for hydrocarbon mixtures from an equation of state with renormalization group theory corrections are performed [2]. We also find several works that presented an efficient and robust algorithm for the calculation of gas-liquid critical point of multi-component petroleum fluids [3]. Other techniques have improved the methods by introducing in their methods new groups which allow describing the various molecular structures and their isomers [4]. A group of researchers presented analytical partial derivative equations required for multi-component critical point calculation [5]. Based on all of the above mentioned, it seems necessary to estimate the critical temperature, which is a representative of maximum temperature at which a gas can be converted into a liquid by an increase in pressure. In addition to that, we find thermal limits CTmax and CTmin are often quantified in terms of critical temperatures using a dynamic measurement [7]. Measurements of thermal limits are increasingly

used to predict organismal responses to climate change and climatic gradients, and empirical estimates of upper and lower thermal limits are now available for a variety of ectotherms [8-10]. Several types of cycloaliphatic hydrocarbons such as cyclopentane, methylcyclopentane and cyclohexane are found extensively in crude oils and some petroleum products [11]. Hydrocarbon mixtures are often supposed as systems which their phase behavior is easy to correlate and predict if their critical points are available [12]. On account of their wide variety of sources, hydrocarbons occur as complex mixtures in the marine environment [13-17]. Quantitative structure- activity relationship (QSAR/ QSPR) are mathematical models designed for the correlation of various types of biological activity, chemical reactivity, equilibrium, physical and physicochemical properties with electronic, steric, hydrophobic and other factors of a molecular structure of a given series of compounds [18]. QSPR uses chemometric methods to describe how a given physicochemical property varies as a function of molecular descriptors describing the chemical structure of the molecule [19]. Chemometrics has provided new insight into the philosophy and theory behind QSPR modeling [20, 21]. QSPR has received significant contributions from various research schools [22, 23]. Application of quantitative structure-property relationship (QSPR) models in prediction and estimation of physical properties of materials is widely developing [24, 25]. In QSPR, advanced mathematical methods (Genetic algorithm, neural networks, and etc.) are used to find a relation between property of interest and the basic molecular properties which are obtained solely from the chemical structure of compounds and called "molecular descriptors" [26]. Also, in drug design and medicinal chemistry, QSAR is one of the treasured implements and most essential areas in chemometric which are comprehensively used [27, 28]. The corresponding experimental data of critical temperature at 1atm were obtained from the literature. It is worth noting that the critical temperature values span between 190.5 and 750 K. Exactly 160 (hydrocarbons, branched alkanes, branched and unbranched alkenes and alkynes) were studied [29]. The structures of the 160 investigated molecules were pre-optimized by means of Molecular Mechanics Force Field, followed by calculations of semi-empirical method [30]. The statistic technique multiple linear regression is used to study the relation between one dependent variable and several independent variables. It is a mathematic technique that minimizes differences between actual and predicted values [31]. The multiple linear regression model (MLR) was generated using the leave-one-out (LOO-CV) cross validation method with the help of Molegro Data Modeller (2009) V. 2.1.0 to predict critical temperature Tc. According to the statistical view point the ratio of the number of samples (n) to the number of descriptors (m) should not be too low. Usually, it is recommended that $n/m \geq 5$ [32]. The PLS method has two objectives: to approximate the matrix X of molecular structure descriptors to the matrix Y of dependent variables and to maximize the correlation between them. The main advantage of this method in comparison to MLR is that interrelated variables can be included in the model [33]. This ought to lead to richer models with better predictive ability [34]. Partial least square analysis was used to check the robustness of the model generated by the multiple least square regression analysis [35]. PLS regression can work well in highdimensional regressions where the number of predictors exceeds the number of observations and set it apart from other predictive methodologies [36]. PLS regression can also be used as a supervised classification method [37]. A recent study developed a fast CV based on the Bouligand influence function (BIF) for kernel-based algorithms [38]. The main objective of this work is the applications and methodologies involved in QSPR focused on obtaining models of the critical temperature, for 160 aliphatic alkanes using calculated constitutional descriptors, atom centered fragments and topological descriptors. In addition to that, a preview and comparison of the models resulting from the statistical calculations using leave-one-out (LOO-CV) cross validation, followed by a discussion of the results.

II. Methodology

II.1. Mathematical QSAR models

QSAR models first summarize the mathematical relationship between chemical structures and physico-chemical property; second predict the property of new chemicals using Multiple Linear and Partial Least Squares Regression analysis. These models are presented by the following form

$$\hat{\mathbf{Y}}_i = \mathbf{b}_0 + \mathbf{b}_1\mathbf{x}_{i,1} + \mathbf{b}_2\mathbf{x}_{i,2} + \ldots\ldots\ldots \mathbf{b}_k\mathbf{x}_{i,k}$$

Where $\hat{\mathbf{Y}}_i$ is the predicted property while **b** values come from statistical software and **x** are the variables (numerical descriptors).

At first, the generated numerical descriptors that encode structural information for the compounds in the data set were calculated. Then, MLR and PLS statistical analysis were used to build the QSPR models. Second comparing these models with each other and third analyzing the explanatory power of these models. The QSPR study was performed in four fundamental stapes:

(1) Selection of data set;
(2) Calculation of molecular descriptors;
(3) Multiple linear (ML) and partial least squares (PLS) regression statistical analysis;
(4) Model validation techniques;

Algerian Journal of Environmental Science and Technology
March edition. Vol.9. Nº1. (2023)
ISSN : 2437-1114
www.aljest.org

The models were evaluated using statistical significance characterization, correlation coefficient R^2, leave-one-out cross validation parameter Q^2, external cross validation Q^2_{ext}, adjusted R^2_{adj}, standard derivation *s* and root mean square error RMSE. The robustness, accuracy and predictive ability of the models were carried out using external cross validation coefficient. Model applicability was further examined by plotting predicted data against experimental data for all the compounds. For example, and not exclusively, the two statistical parameters, Q^2 and RMSE, can be calculated according to the following equations**:**

$$Q^2 = 1 - \frac{\sum (Yi - Yipred)^2}{\sum (Yi - Ymean)^2}$$

Where $Y_i$is the i^{th} experimental T_c value, $Y_{ipred}$is the i^{th} predicted T_c,$Y_{mean}$is the mean of the experimental T_c. The summation is over all patterns in the analysed data set.

$$RMSE = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} (p_{exp} - p_{pred})^2}$$

Where **n** = number of compounds, $\mathbf{p_{exp}}$ is the experimental value, $\mathbf{p_{pred}}$ is the predicted value

II.2. Structure and molecular descriptors

The selected molecules were pre-optimized by the Molecular Mechanics (MM+) in the Hyper Chem 7,5 program. The minimized structures were refined using the semi-empirical PM3 method at a restricted Hartree Fock level with no configuration interaction, applying a gradient norm limit of 0.01 kcal Å -1 mol^{-1} as a stopping criterion. Then the structures were used as input for the generation of molecular descriptors. The total number obtained after the calculation is122 (2D) descriptors from three different blocks: 30 constitutional descriptors, 13 atom centered fragments and 79 topological descriptors were used to characterize the set of 160 aliphatic alkanes. The descriptors calculated using the DRAGON (version 5.4) software**.** These descriptors were analyzed to check and remove constant or near-constant variables. The remaining descriptors were used to build the *X*-matrix in the MLR and PLS regression analysis, as follows: 14 constitutional descriptors, 9atomcentered fragments and 35 topological descriptors.

II.3. Methodology QSPR

The proposed methodology QSPR is illustrated in Figure 1. Developing QSAR models starts with the collection of the experimental values for proper of interest, taking into consideration the source and quality of this data. This is in order to avoid everything that has a negative impact on the validity and robustness of the model.

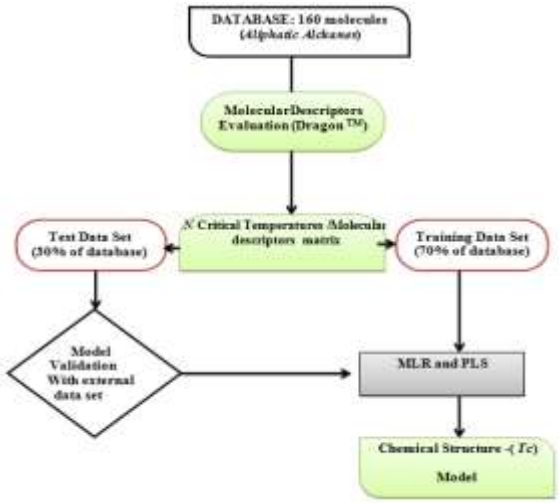


Figure 1. *Steps of QSPR proposed methodology*

II.4. Training and test set generation.

In a QSAR / QSPR study, generally, the quality of a model is expressed by its fitting ability, and prediction ability, and of these the prediction ability is the more important. In order to build and test the model, a several procedures can be adopted for the selection of the training and test sets, whereas the training/test sets were built taking randomly 30 % of each cluster as the test set, while the remaining 70 % were used as the training set.

III. Method and discussions

In order to build and test the models, a data set of 160 compounds was separated into a training set of 111 compounds, which were used to build the model and a test set of 49 compounds, which were applied to test the built models. These compounds were quantified with the help of molecular descriptors, calculated with the DRAGON software. All regression calculations were performed with the aid of the Molegro software package. Based on the descriptors matrices of MLR and PLS for three molecular blocks constitutional descriptors, atom centered fragments andtopologicaldescriptors, one can calculate their

QSPR analyzes,by searching for the most significant descriptors which have a strong degree of correlation with the studied propertyT_c.

III.1. Multiple linear regressions

A. Constitutional descriptors block

Figure 2 shows the experimental versus predicted values by MLR models. It shows also the critical temperature **Tc** (exp) versus predicted **Tc** (pred) in the constitutional block, the atom centered fragment block and the topological description block. To establish quantitative relationships between critical temperature **Tc** and selected descriptors, our array data were subjected to a multiple linear regression. Many attempts have been made to develop a relationship with the indicator variable of critical temperature **Tc**, the best relationship obtained by this method is only one corresponding to the linear combination of several descriptors featured in the following equation (3) and their interpretations are shown in Table 1.

Tc = 5.14185 + 292.314 * Sv - 254.725 * Sp - 71.2192 * Ms + 8.52197 * NBO - 4.0058 * SCBO - 11.3033 * RBF - 1.48307 * nC (3)

B. Atom centered fragments block

The following equation represents the MLR model of the critical temperature T**c** for the atom centered fragments block. In this calculated model we find several molecular descriptors included the following equation (4) and their abbreviations are shown in Table 1. It should be noted that the symbols associated with these descriptors are: R: represents any group linked through carbon. X: represents any electronegative atom (O, N, S, P, Se, halogens). #: represents a triple bond .a the superscript represents the formal oxidation number.

Tc = 4.29504 -238.504 * "C-001" - 138.657 * "C-002" - 51.7308 * "C-003" + 60.0144 * "C-004" - 35.1493 * "C-015" + 22.6861 * "C-016" + 78.5786 * "C-022" + 80.2691 * "H-046" + 2.4934 * "H-047" (4)

C. Topological descriptors block

The MLR equation of topological descriptors bloc is composed by the several descriptors featured in the following equation (5) and their meanings are shown in Table 1.

Tc = 3.72571 + 1.77888 * ZM1V + 1.69172 * SMTI + 0.0571746 * SMTIV + 5.53309 * D/D - 0.719466 * S0K + 0.89795 * CSI - 0.543406 * CENT - 0.176002 * HyDp - 0.937503 * Wap - 5.61055 * QW (5)

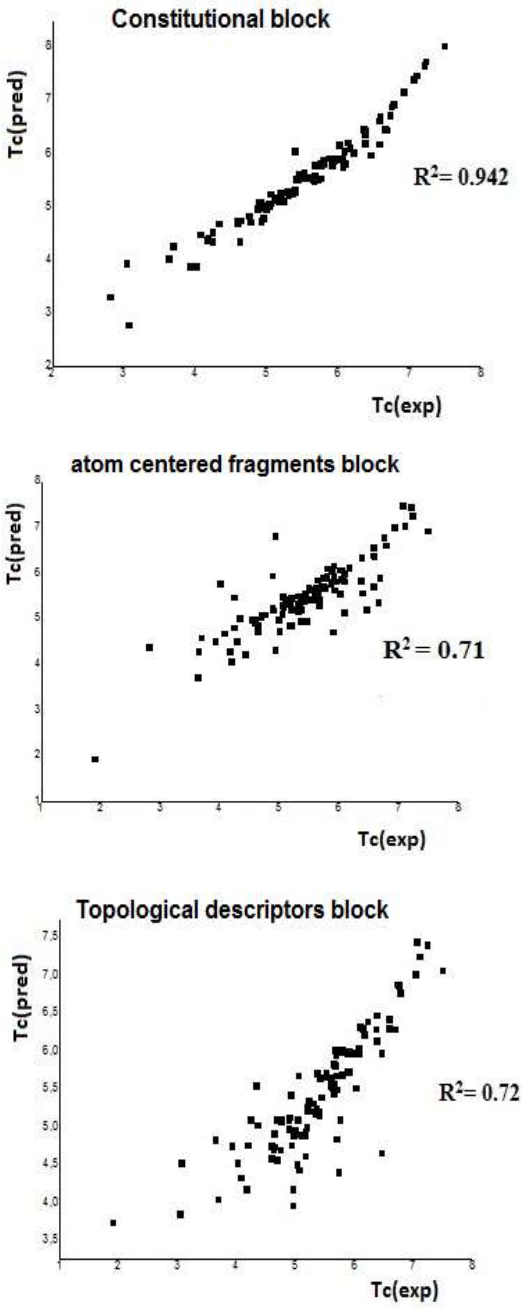


Figure 2. *Plot of T_c Experimental vs. T_c predicted values by MLR models*

III.2.Partial least square regressions

Figure 2 shows the experimental versus predicted values by PLS models. It shows also the criticaltemperature T_c (exp) versus predicted T_c (pred) in the constitutional block, the atom centered fragment block and the topological description block.

A. Constitutional descriptors block

The PLS equation of constitutional descriptors bloc is composed by the several descriptors featured in

the following equation (6) and their meanings are shown in Table 1.

T_c =7987245.5854 -1246.9860 Sp+3688.1330Ms-190.6498nat+21867947.05274 nSk -588.5911 Nbo+190.6494 Nh (6)

B. Atom centered fragments block

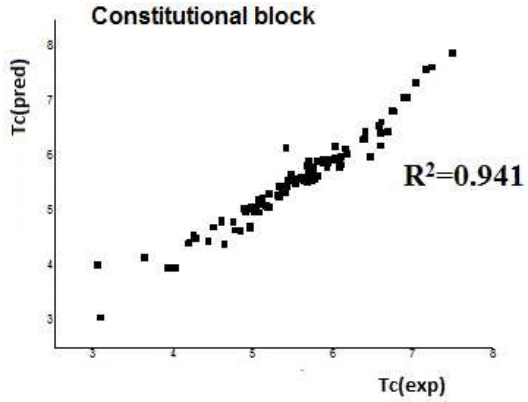
The PLS model equation of atom centered fragments block is composed by the several descriptors featured in the following equation (7) and their abbreviations are shown in Table 1.

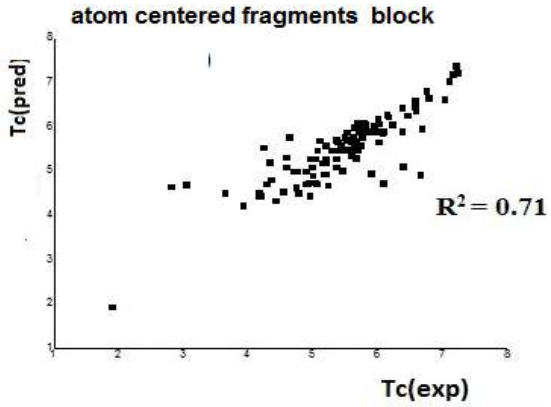
T_c = -32.5633 -279.7931*"C-001"-166.4402*"C-002 -71.7958*"C-003" +51.8108*"C-004" -25.2363* "C-015+4.9583*"C-016+92.7512* "H-046 (7)

C. Topological descriptors block

The PLS model equation of Topological descriptors block is composed by the several descriptors featured in the following equation (8) and their interpretations are shown in Table 1.

T_c =- 12.70139+ 3.8384*ZM1V+0.08276*SMTI-0.26197*SMTIV+2.6429*D/D +1.1507*S0K+1.0118*CSI-0.3254*CENT (8)





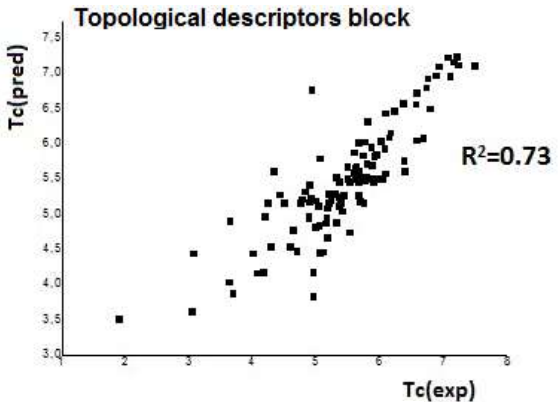


Figure 3. *Plot of T_c Experimental vs. T_c predicted values by PLSmodels*

Table 1. *The Descriptors found in various calculated models*

Category of descriptors	Name of the descriptors
Constitutional descriptors	The sum of atomic van der Waals volumes(**Sv**) The sum of atomic polarizabilities (**Sp**) The number of atoms(**Ms**) Number of non-H bonds(**Nbo**) Sum of conventional bond orders (H-depleted) (**SCBO**) Rotatable bond fraction(**RBF**) Number of Carbon atoms(**nc**) Number of atoms(**nAT**) Number of non-H atoms(**nSK**) Number of Hydrogen atoms(**nH**)
Atom-centred fragments descriptors	Described byCH_3R / CH_4(**C-001**) Described byCH_2R_2, (**C-002**) MeansCHR_3(**C-003**) MeansCHR_4(**C-004**) Means =CH2(**C-0015**) Means =CHR(**C-0016**) which means **#CR / R=C=R**(**C-0022**) MeansH^a attached to C^0(sp3) no X attached to next C(**H-046**) MeansH^a attached to C^1(sp3) / C^0(sp2) (**H-047**
Topological descriptors	The sum of atomic van der Waals volumes(**ZM1V**) The sum of atomic polarizabilities (**SMTI**) Number of non-H bonds(**D/D**) Sum of conventional bond orders (H-depleted) (**S0K**) Rotatable bond fraction(**CSI**) Hyper-distance-path index(**HyDp**) All-path Wiener index(**Wap**) Quasi-Wiener index (Kirchhoff number) (**QW**)

III.3. Discussions

The statistical worthiness of the constitutional descriptors block developed model was evaluated in terms of square of the correlation coefficient, where R^2 (MLR=0.942 and PLS=941) values explain 94 % variance in critical temperaturesT_c, which indicates a measure of good fit by the regression equations. We notice from the comparison Table 2 that the slight difference in (MLR) R^2 (0.942) and Q^2_{Loo} (0.834) is equal to 0.108 it can be considered acceptable, these values implies high prognostic ability of themodel. Similarly in the (PLS) analysis, there is a very small difference in R^2 (0.941) and Q^2_{Loo}(0.932) values, further ascertains the robustness of the model.

Table 2. *Results comparison Table*

Block	$N_{training}$	N_{test}	R^2	Q^2_{Loo}	Q^2_{ext}	R^2_{adj}	RMSE	s
MLR model								
Constitutional	111	49	0.942	0.83	0.794	0.938	0.22	0.88
Atom centred fragments	111	49	0.71	0.56	0.62	0.68	0.49	0.77
Topologic Descriptors	111	49	0.72	0.63	0.51	0.70	0.47	0.76
PLS model								
Constitutional	111	49	0.941	0.932	0.80		0.20	0.797
Atom centred fragments	111	49	0.71	0.67	0.55		0.47	0.74
Topologic Descriptors	111	49	0.73	0.695	0.41		0.48	0.80

It is worth noting that the PLS model is completely satisfactory in the fitting and has high predictive power compared to the MLR model. The LOO-CV (leave-one-out) cross-validation highlights that the model is stable, not obtained by chance, in fact the difference between R^2 and Q^2_{Loo} is small: 0,9 %. However MLR is high is estimated: 10%. It is not limited to that only, but we also see that the number of variables (descriptors) in the PLS model equation (6) is less than the number in the MLR model equation (3). As for the value of RMSE=0.20.The final model has the highest correlation coefficient (0.94), confirming the robustness of the model and also the PLS model was cross validated and the Q^2_{Loo} value of (0.932) depicted the strength of the model, it satisfy the conditions: (1) R^2> 0.7, (2) Q^2_{Loo}> 0.6, (3) difference between R^2 and Q^2_{Loo} smaller than 0.1[39, 40]. Based on all of the above, it can be said that the MLR model is less stable than PLS model.The value'**s**' is the standard (deviation) error of the regression model, and it should be low for better QSPR model generation; this is an approximation of how precisely the model will predict unknown 'Y' values. The value of '**s**' for the best PLS model is 0.797 while for MLR is 0.88. It signifies that regression with an 's' value of 0.8 should predict Y values with a standard error of

0.8 units. The large value of the standard deviation is due to the difference between the experimental values of $\mathbf{T_c}$, which span between 190.5 and 750 K. As for the models atom centered fragments and Topological descriptors blocks, for both MLR and PLS regression, we find that the validation criteria, leave-one-out cross-validated Q^2_{Loo}. According to Hawkins et al., a valid statistical model should have high Q^2 value (Q^2_{Loo}>0.5) and is evidence of the high predictive ability of the model [39, 40]. High or acceptable values of the two parameters, Q^2_{Loo} and R^2, may be obtained as long as a moderate overall correlation is maintained between the observed and predicted T_c values even if there is a considerable difference between them. Depending also on the conditions[39, 40], as it is noticed that it is achieved in the mentioned models except the MLR model of atom centered fragments block, therefore, it can be said that the models are acceptable considering the values of the traditional parameters (Q^2 and R^2).

IV. Abbreviation list

QSPR. - quantitative structure-property relationship
QSAR. - quantitative structure-activity relationship
MLR t. - Multiple Linear Regression apt. - apartment
PLS. - Partial Least Squares
T_c. - critical temperature
R^2. - squared correlation coefficients
CT. - thermal limits
LOO-CV. - leave-one-out cross validation
BIF. - Bouligand influence function
Q^2. - leave-one-out cross validation parameter
Q^2_{ext}. - external cross validation
R^2_{adj}. - adjusted squared correlation coefficients
s. - standard derivation
RMSE. -root mean square error

V. Conclusion

In this study, we have compared the performance of, MLR and PLS in (Aliphatic Alkanes) QSPR study, using the selected descriptors, which can be easily generated from the Dragon software. The obtained results showed that:

- The two constructed of PLS and MLR constitutional block models have a good reliability (R^2train > 0.94, Q^2train > 0.834) and predictability (Q^2_{ext}> 0.79);
- The overall performance of prediction was found to be around 94% in case of PLS and MLR;
- The difference between R^2 and Q^2_{Loo} is much less than 0.1 and that's what makes us say that PLS is significant predictive power and reliability as compare to MLR technique;
- Four other models have also been built, represented by PLS and MLR models of atom cent ered fragments and Topological descriptors blocks that can be considered to have acceptable predictive power.

VI. References

1. Peisheng, M.A.; Jin, G.A.O.; and Shuqian, XIA. Determination of the Critical Temperature and Critical Pressure of Five Compounds. *Chinese J.Chem. Eng*. 10 (2002) 473-375.
2. Castier, M.; Sandler, S.I. Critical points with the Wong-Sandler mixing rule- II. Calculations with a modified Peng-Robinson equation of state. *Chemical Engineering Science*. 52 (1977) 3579-3588.
3. Jiang, J.; Prausnitz, J.M. Critical temperatures and pressures for hydrocarbon mixtures from an equation of state with renormalization-group theory corrections. *Fluid Phase Equilibria. J.M.* (2000) 169.
4. Marrero, J.; Gani, R. Group-contribution based estimation of pure component properties. *Fluid Phase Equilibria*. (2001) 183–208.
5. Chaikunchuensakun, S.; Tanthapanichakoon, W. Critical points calculation with a cubic equation of state and excess free energy mixing rules. *Fluid Phase Equilibria*. (2003) 113-129.
6. Hoteit, H.; Santiso, E.; Firoozabadi, A. An efficient and robust algorithm for the calculation of gas-liquid critical point of multicomponent petroleum fluids. *Fluid Phase Equilibria*. 241(2006) 186–195.
7. Kingsolver, J. G.; Umbanhowar, J. The analysis and interpretation of critical temperatures. *Journal of Experimenta Biology*. (2018) 221. doi:10.1242/jeb.167858
8. Chowen, S.L.; Terblache, J.S. Physiological Diversity in Insects:Ecological and Evolutionary Contexts.*Adv Insect Physiol*. 33 (2007) 50-152.
9. Sunday, J.M. Bates, A.E .; Dulvy, N. Global analysis of thermal tolerance and latitude in ectotherms. *Proc. R. Soc*. 278(2011) 1823-1830.
10. Buckleym, L.B.; Huey, R.B. Temperature extremes: geographic patterns, recent changes and implications for organismal vulnerabilities .Glob. *Change Biol*. 22(2011) 3829-3842.
11. Comandini, A.; Dubois, T.; Abid, S.; Chaumeix, N. Comparative study on cyclohexane and decalin oxidation. *Energy & Fuels*. 28 (1) (2013) 714–724.
12. Hosseini, S.M.; Fazlali1, A.R.; Zahed, G.R. Critical properties of hydrocarbon mixtures by artificial neural network and peng-robinson EOS Arpn. *Journal of Engineering and Applied Sciences*, 3 (6) (2008) 1819-6608.
13. Bouloubassi, I.; Lipiatou, E.; Saliot, A.; Tolosa, I.; Bayona, J. M.; Albaiges, J. Carbon sources and cycle in the western Mediterranean – The use of molecular markers to determine the origin of organic matter, *Deep-Sea Res*. 44 (1997) 781–799.
14. Gogou, A.; Bouloubassi, I.; Stephanou, E.G. Marine organic geochemistry of the Eastern Mediterranean: 1. Aliphatic and polyaromatic hydrocarbons in Cretan Sea surficial sediments, *Marine Chemistry*. 68(2000) 265–282.
15. Prahl, F.G.; Carpenter, R. Hydrocarbons in Washington coastal sediments, *Estuar. Coast. Shelf S.* 18 (1984) 703–720.
16. Tolosa, I.; Bayona, J. M.; Albaiges, J. Aliphatic and polycyclic aromatic hydrocarbons and sulfur/oxygen derivatives in Northwestern Mediterranean sediments: spatial and temporal variability fluxes and

budgets. *Environ. Sci. Technol*. 30 (1996) 2495–2503.
17. Yunker, M.B.; Macdonald, R.W.; Vingarzan, R.; Mitchell, R. H.; Goyette, D.; Sylvestre, S. PAHs in the Fraser River basin: A critical appraisal of PAH ratios as indicators of PAH source and composition.*Org. Geochem.* 33 (2002) 489–515.
18. Shafiei, F. Relationship between Topological Indices and Thermodynamic Properties and of the Monocarboxylic Acids Applications in QSPR.Iranian *Journal of Mathematical Chemistry*. 6 (2015) 15-28.
19. Fengping, L.; Chenzhong, C.; Bin, C. A Quantitative Structure-Property Relationship (QSPR) Study of Aliphatic Alcohols by the Method of Dividing the Molecular Structure into Substructure. *Int. J. Mol. Sci*. 12(2011) 2448-2462.
20. Katritzky, A.R.; Petrukhin, R.; Tatham, D. Interpretation of quantitative structure-property and activity relationships. *J. Chem. Inf. Comput. Sci.* 41(2001)679–685.
21. Katritzky, A.R.; Dobchev, D.A.; Slavov, S.; Karelson, M. Legitimate utilization of large descriptor pools for QSPR/QSAR models. *J. Chem. Inf. Model*. 48 (2008) 2207–2213.
22. Delgrado, E.J.; Alderete, J.B.; Gonzalo, A.J. A simple QSPR model for predicting soil sorption coefficients of polar and nonpolar organic compounds from molecular formula. *J. Chem. Inf. Comput. Sci*. 43 (2003) 1928–1932.
23. Laura, D.H.; David, S.P.; Florian, N.; John, B.O. Why are some properties more difficult to predict than others? A study of QSPR models of solubility, melting point, and Log P. *J. Chem. Inf. Model*. 48 (2008) 220–232.
24. Katritzky, A.R.; Fara, D.C. How chemical structure determines physical, chemical, and technological properties: an overview illustrating the potential of quantitative structure-property relationships for fuels science. *Energy & Fuels*. 19 (2005) 922–935.
25. Taskinen, J.; Yliruusi, J. Prediction of physicochemical properties based on neural network modeling. Adv. *Drug Delivery Rev*. 55 (9) (2003) 1163-1183.
26. Vatani, A.; Mehrpooya, M.; Gharagheizi,F. Prediction of Standard Enthalpy of Formation by a QSPR Model, *Int. J. Mol. Sci*. 8(2007) 407-432.
27. Prachayasittikul, V.; Pingaew, R.; Worachartcheewan, A.; Nantasenamat, C. Prachayasittikul, S. Ruchirawat, S. Prachayasittikul, Virapong. Synthesis, anticancer activity and QSAR study of 1,4-naphthoquinone derivatives. *Eur. J. Med. Chem*. 84 (2014) 247–263.
28. Prachayasittikul, V.; Pingew, R.; Anuwongcharoen, N.; Worachartcheewan, A.; Nantasenamat, C.; Prachayasittikul, S.; Prachayasittikul, V. Discovery of novel 1, 2,3-triazole derivatives as anticancer agents using QSAR and in silico structural modification. *Springer Plus* 4 (1) (2015) .
29. Reid, R.C.; Prausnitz, J.M.; Poling, B.E. The Properties of Gases & liquids, Fourth Edition, *McGraw-Hill Book Company*, New York, (1987).
30. Araújo, P. H. F. ; Ramos, R.S. ; da Cruz, J. N. ; Silva, S. G.; Ferreira, E. F. B. ; de Lima, L. R.; Macêdo, W. J. C. ; Espejo-Román, José. M.; Campos, J.M.; Santos, C.B.R. Identification of Potential COX-2 Inhibitors for the Treatment of Inflammatory Diseases Using Molecular Modeling Approaches. *Molecules* .25 (8) (2020) 4183.
31. Mbarki, S.; Dguigui,K .; El Hallaoui, M. Construction of 3D-QSAR models to predict antiamoebic activities of pyrazoline and dioxazoles derivatives . *J. Mater. Environ. Sci*. 2 (1) (2011) 61-70.
32. Xu, J.; Zhang, H.; Wang, L.; Liang, G.; Wang, L. ; Shen, X.; Xu, W. *Spectrochimica Acta Part A*. 76 (2010) 239.
33. Ivan, D.; Crisan, L.; Funar-Timofei, S.; Mracec, M. A quantitative structure–activity relationships study for the anti-HIV-1 activities of 1-[(2-hydroxyethoxy) methyl]-6- -(phenylthio)thymine derivatives using the multiple linear regression and partial least squares methodologies. *J. Serb. Chem. Soc*. 78 (4) (2013) 495–506.
34. De Fortuny, E.J.; Martens, D.; Provost, F. Predictive modeling with big data: Is Bigger Really Better? *Mary Ann Liebert, Inc*. 1(2013).
35. Kesar, S .; Paliwal, S.K.; Mishra, P.; Chauhan, M. Quantitative Structure–Activity Relationship Analysis of Selective Rho Kinase Inhibitors as Neuro-regenerator Agents *Turk J Pharm Sci*. 16(2) (2019)141-154.
36. Dennis Cook, R. Envelopes: A new chapter in partial least squares regression *journal of Chemometrics First published*: 15 September (2020) .
37. Kvalheim, T.R. ; Kvalheim,O. M. Multivariate data analysis in pharmaceutics: A tutorial review. *International Journal of Pharmaceutics*. 417 (1-2) (2011) 280-290 .
38. Liu, Y.; Liao, S.; Jiang, S.; Ding, L.; Lin, H.; Wang, W. Fast Cross-Validation for Kernel-based Algorithms. *IEEE Trans Pattern Anal Mach Intell*.42(2019) 1083–1096.
39. Chirico, N.; Gramatica, P. Real external predictivity of QSAR models: How to evaluate it? Comparison of different validation criteria and proposal of using the concordance correlation coefficient. *J. Chem. Inf. Model*. 51 (2011) 2320–2335.
40. Hawkins, D.M.; Basak, S.C.; Mills, D. Assessing model fit by cross-validation, *J. Chem. Inf. Comput*. 43 (2) 2003 579–586.

Please cite this Article as:
Souyei B., Kertiou N., Hadj Seyd A., Labbi Y., Khechekhouche A., Critical Temperatures Prediction of Organic Compounds (Aliphatic Alkanes) Using QSPR Approach, ***Algerian J. Env. Sc. Technology***, ***9*:1 (2023) 2882-2889**